What you always wanted to know about mouse gene and strain nomenclature...

Debbie Krupke

Standard Nomenclature

- Advantages
  - Unique identifier
  - Conveys key information
- Uses
  - Publications
  - Ordering strains
  - Database integration

Mouse Nomenclature

- http://www.informatics.jax.org/mgihome/nomen/
  - Rules for Nomenclature of Genes, Genetic Markers, Alleles, and Mutations in Mouse and Rat
  - Quick Guide to Nomenclature for Alleles and Mutations
  - Mutation Submission Form
  - Rules for Nomenclature of Mouse and Rat Strains
  - Mouse Strain Registry Form
  - New Nomenclature for Strain 129 Mice

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Gene Nomenclature

- Symbols
  - A mouse gene symbol is italicized and begins with a capital letter unless the gene is known only by its phenotype and the phenotype is recessive.
  - A human gene symbol is italicized and entirely upper case.
    - Exception: "Because they may be difficult to read, depending on the browser, gene symbols are frequently not italicized when posted to a web page."
- Names
  - Both mouse and human gene names are written in roman letters and entirely lower case (unless they include a proper name).

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Spontaneous Mutations

- When a novel spontaneous mutation is discovered, it is assigned a descriptive name & symbol (*e.g.*, *scb*, scabby).
- If a named mutation is found to be allelic with a previously named mutation, its symbol is superscripted to the symbol for the earlier mutation (*e.g.*, hr^{rh}).
- Once the gene underlying a mutation has been identified the former descriptive gene symbol is superscripted to the new gene symbol to form an allele symbol (*e.g.*, *ob* -> Lep^{ob}). If the previous symbol already contained a superscript the former parts of the symbol are superscripted and separated by a hyphen (*e.g.*, Mo^{br} -> $Atp7a^{Mo\text{-}br}$).

Spontaneous Mutations

- *Lep*^{ob} = allele symbol for the obese allele of the leptin gene
  - *Lep* = gene symbol
  - leptin = gene name
  - *ob* refers to an allele
  - obese = allele name

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Random Mutagenesis

Fech^{m1Pas}

Fech = symbol of the affected gene

m1 = mutation 1 of this gene in this lab

Pas = laboratory code for the Institut Pasteur

These types of mutations often result from chemical mutagenesis experiments.

Laboratory Registration Codes

- A series of (usually three to four) letters that identifies a particular institute, laboratory, or investigator that produced and/or holds stocks of a strain.
  - First letter upper case
  - Subsequent letters lower case
- Lab Codes are assigned by the Institute for Laboratory Animal Research (ILAR)
  - http://dels.nas.edu/ilar/codes.asp?id=codes

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Targeted Mutations

Kras2^{tm1Tyj}

Kras2 = symbol of the targeted gene

tm1 = targeted mutation 1 of this gene in this lab

Tyj = laboratory code for Tyler Jacks

Floxed/Recombined Alleles

- Floxed alleles are named just as other targeted mutant alleles
  - *Brca1*^{tm2Cxd}
    - Floxed allele of *Brca1* generated in the lab of Chu-Xia Deng. This is the second targeted allele of *Brca1* made in his lab.
- Once the allele has been recombined in germ cells ".1" is appended to the targeted mutation number.
  - *Brca1*$^{tm2.1Cxd}$
    - This allele is the result of Cre recombination of *Brca1*^{tm2Cxd} in the germline.

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Transgene Nomenclature

Tg(Ins2-cre)25Mgn

Tg = transgenic

(Ins2-cre) = insert designation

Ins2 = regulatory component; rat insulin II promoter

cre = expressed component; bacterial cre recombinase

Exception: Insert designations from alleles named before the rule was amended may not follow the (PromoterSymbol-GeneSymbol) convention.

25 = founder line number

Mgn = laboratory code for Mark A. Magnuson

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Inbred Strains

- A strain is inbred if it has been bred sister x brother for 20 or more consecutive generations (F ≥ 20).
- An inbred strain is designated by a unique, brief symbol made up of upper case roman letters or a combination of letters and numbers beginning with a letter.
- Some traditional strains do not follow this convention (*e.g.*, 129P1/J).

Revised 129 Nomenclature

- Most ES cell lines derived from 129 strains
- Great genetic variability among 129 substrains
- 129 substrain names reflect parental strain
  - Parental Steel Teratoma X - (contaminated)

129P2/OlaHsd	129S1/SvImJ	129T2/SvEms	129X1/SvJ

- http://jaxmice.jax.org/html/nomenclature/nomen_129.shtml

FVB/N

- FVB = parent strain designation
- N = laboratory code for NIH

C57BL/6J

- C57BL = parent strain designation
- 6J = substrain designation
  - 6 = line number
  - J = laboratory code for The Jackson Laboratory

Substrains

- A strain gives rise to a substrain, or a substrain to further strains, due to:
  - Separation after F20 but before F40
  - Separation from the original strain for >20 generations
  - Detection of genetic differences
- A substrain is identified by appending to the root strain name a "/" followed by the Lab Code of the holder.
- Further substrains are designated by addition or the Lab Codes of subsequent holders of the strain, without another forward slash.

C3H/HeOuJ

- C3H = parent strain designation
- He = laboratory code for Walter E. Heston
- Ou = laboratory code for Henry C. Outzen
- J = laboratory code for The Jackson Laboratory

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

F1 Hybrid Mice

- Progeny of two inbred strains crossed in the same direction (F x M)
- Genetically identical (except XX, XY differences)
- Can be designated using upper case abbreviations of the parent strains followed by "F1"
- Maternal strain listed first

F1 Hybrid Nomenclature

- B6D2F1 or (C57BL/6J x DBA/2J)F1
  - Female progenitor: C57BL/6J
  - Male progenitor: DBA/2J
- CSJLF1 or (BALB/cJ x SJL/J)F1
  - Female progenitor: BALB/cJ
  - Male progenitor: SJL/J

Standard Abbreviations

129P3/J = 129P3

129S1/SvImJ = 129S1

A/HeJ = AHe

A/J = A

AKR/J = AK

BALB/cByJ = CBy

BALB/cJ = C

C57BL = B

C57BL/6J = B6

C57BL/6JEi = B6Ei

C57BL/10J = B10

C57BR/cdJ = BR

C57L = L

CBA/J = CBA

C3H/HeJ = C3

C3HeB/FeJ = C3Fe

DBA/1J = D1

DBA/2J = D2

NZB/BlNJ = NZB

NZW/LacJ = NZW

SJL/J = SJL or J

SWR/J = SW

Mouse Nomenclature

- Genes
- Alleles
  - Spontaneous Mutations
  - Induced Mutations
    - Random Mutagenesis
    - Targeted Mutations
    - Transgenes
- Strains
  - Inbred Strains
  - Hybrid Mice
  - Mutant Mice

Genetic Background

- Inbred
  - Mice were generated on and maintained on the same inbred genetic background.

    C57BL/6-*Apc*^{Min}

    The "Min" mutation arose in a C57BL/6 mouse and mice have been maintained on a pure C57BL/6 genetic background.
  - Mice that were generated on one genetic background and have subsequently been backcrossed to another genetic background DO NOT fall into this category (no matter how long the backcrossing has gone on).
- F1 Hybrid
  - Mutant mice generated on an inbred background and crossed ONCE to another inbred background.

    AKB6F1-*Apc*^{Min}

    Female AKR mouse crossed to a male C57BL/6-*Apc*^{Min} mouse. This can also be written as (AKR/J x C57BL/6J)F1-*Apc*^{Min}.

Genetic Background

- Congenic
  - Produced by repeated backcrosses to an inbred strain, with selection for a particular marker from the donor strain. Minimum of 10 backcross generations (N10) needed for full congenic.
- Incipient Congenic
  - Backcross generations N5 - N9

Congenic and incipient congenic strains are designated by a period between the recipient and donor strain abbreviations (*e.g.* B6.129P2-*Apoe*^{tm1Unc}).

Genetic Background

- Mixed background (two strains)
  - A semicolon indicates that the genetic background of a stock is a mixture of the strains whose abbreviations are separated by the semicolon, with no implication of their relative proportions.
- B6;129S7-*Acvr2*^{tm1Zuk}
  - A mixed strain derived from C57BL/6J and a 129 ES cell line carrying a targeted mutation of the *Acvr2* gene; continued sister x brother mating leads to a new inbred strain.

Genetic Background

- STOCK
  - A strain, incipient or inbred, whose genome - excluding a congenic segment - is derived from more than two inbred progenitor strains or includes genetic contribution from an unknown or outbred source is considered a "mixed" stock or strain.
  - The strain is designated "STOCK" followed by a space and the mutation(s) or chromosome anomaly it carries.
  - STOCK *Egfr*^{tm1Mag}

Genetic Background

inbred	Strain generated and maintained on the same inbred genetic background.
F1 hybrid	Strain generated on an inbred genetic background and crossed once to another inbred background.
period .	Congenic (backcrossed to N10 or greater) Incipient congenic (backcrossed to N5-N9)
semicolon ;	The genetic background is some mixture of the strains separated by the semicolon, with no implication of proportions thereof.
STOCK	Mixed genetic background derived from more than two progenitor strains or having genetic contribution from an unknown or outbred source

Mutant Mice Nomenclature

Genetic background + Allele(s)

Strains with Spontaneous Mutations

Genetic background + Allele(s)

C57BL/6J-*Kit*$^{W\text{-}v}$

B6.C3-*Kit*$^{W\text{-}44J}$

B6;AKR-*Lxl2*

STOCK *Kit*$^{Sl\text{-}16J}$

Strains with Random Induced Mutations

Genetic background + Allele(s)

C57BL/6J-*Apc*^{Min}/J

AKB6F1-*Apc*^{Min}

AK.B6-*Apc*^{Min}

B6;AK-*Apc*^{Min}

STOCK *Apc*^{Min}

Strains with Targeted Mutations

Genetic background + Allele(s)

129S6/SvEv-*Mos*^{tm1Ev}

B10.129P2(B6)-*Il10*^{tm1Cgn}/J

B6;129S7-*Rb1*^{tm1Brd} *Trp53*^{tm1Brd}

STOCK *Mos*^{tm1Ev}

Strains with Transgenes

Genetic background + Allele(s)

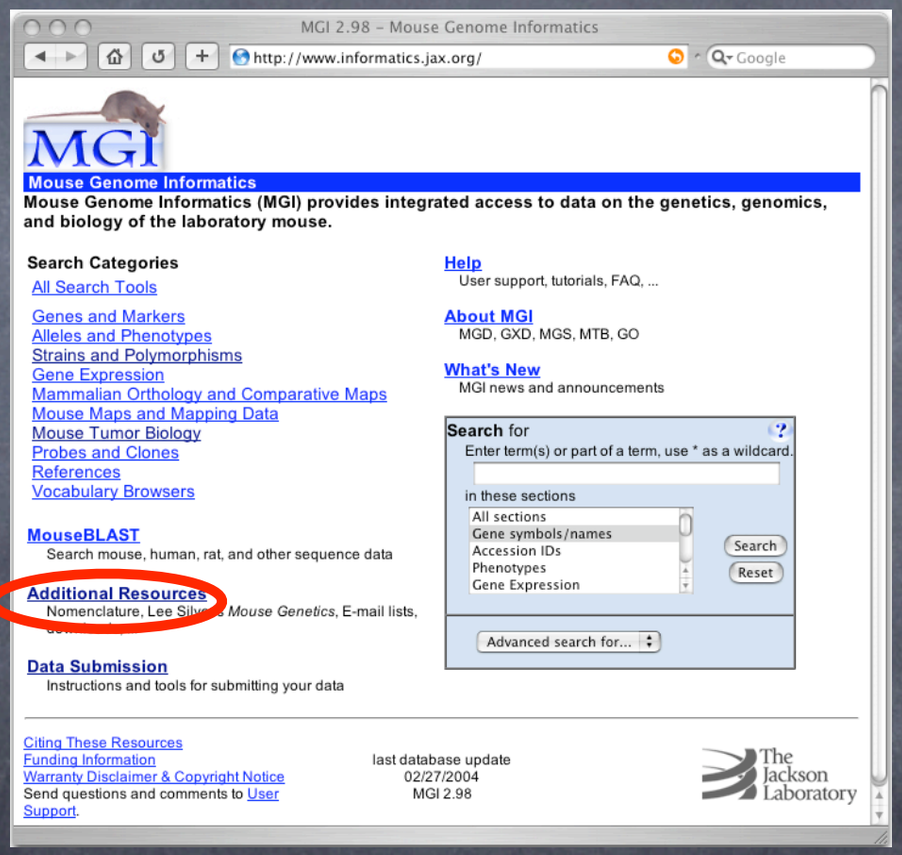
FVB/N-Tg(MMTVPyVT)634Mul

(NOD/LtJ x FVB/N)F1-Tg(MMTVPyVT)634Mul

FVB;129-Tg(MMTVPyVT)634Mul

STOCK Tg(MMTVPyVT)634Mul

Mouse Nomenclature Questions?



http://www.informatics.jax.org/mgihome/nomen/

nomen@informatics.jax.org